REPORT RESUMES

ED 016 282 24 EA 001 020
COGNITIVE STRUCTURES AND EDUCATIONAL EVALUATION.
BY- SKAGER, R.W. BROADBENT, L.A.
CALIFORNIA UNIV., LOS ANGELES
REPORT NUMBER CSEIP-OR-5 PUB DATE APR 67
REPORT NUMBER BR-6-1646
CONTRACT OEC-4-6-061646-1909-1
EDRS PRICE MF-$0.50 HC-$3.56 87P.

DESCRIPTORS- *COGNITIVE DEVELOPMENT, *COGNITIVE MEASUREMENT, COGNITIVE PROCESSES, COGNITIVE TESTS, INSTRUCTIONAL PROGRAMS PROBLEM SOLVING, *CONCEPT FORMATION, *EVALUATION, *LEARNING PROCESSES, CURRICULUM EVALUATION, MEASUREMENT TECHNIQUES, ACHIEVEMENT TESTS, PIAGETIAN THEORY, LOS ANGELES,

OVER 40 ARTICLES AND STUDIES RELATED TO THE PIAGETIAN THEORY OF COGNITIVE DEVELOPMENT ARE CRITICALLY REVIEWED AS A BASIS FOR SUGGESTING APPLICATIONS OF THE THEORY TO EDUCATIONAL EVALUATION. PIAGET'S CONCEPTS ARE ANALYZED WITHIN TWO BROAD CATEGORIES, THE STAGE OF CONCRETE OPERATIONS, FROM APPROXIMATELY SEVEN TO 11 YEARS OF AGE, AND THE STAGE OF FORMAL OPERATIONS, FROM APPROXIMATELY 12 YEARS OF AGE TO ADULTHOOD. DURING THE STAGE OF CONCRETE OPERATIONS, THE CHILD DEVELOPS REPRESENTATIONS OF THE PHYSICAL WORLD AND THE ABILITY TO UTILIZE THE LOGICS OF CLASS INCLUSION AND SERIAL ORDERING. COGNITIVE FUNCTIONING AT THIS STAGE IS CONFINED MAINLY TO DEALING WITH THE ACTUAL OR PHENOMENAL RATHER THAN THE POSSIBLE OR HYPOTHETICAL. DURING THE STAGE OF FORMAL OPERATIONS, THE CHILD BECOMES ABLE TO DIFFERENTIATE VARIABLES IN A PROBLEM-SOLVING SITUATION AND TO ISOLATE CRITICAL MANIPULATIONS AT A SYMBOLIC LEVEL. FURTHER RESEARCH IS NEEDED, UTILIZING PIAGETIAN TASK MEASUREMENT CONCEPTS, TO DETERMINE CLEARLY THE EFFECTS OF CULTURE AND TRAINING ON COGNITIVE DEVELOPMENT AS WELL AS THE PART PLAYED IN COGNITIVE DEVELOPMENT BY SUBJECT CHARACTERISTICS OTHER THAN CHRONOLOGICAL AGE. PIAGETIAN TASK MEASUREMENT CONCEPTS APPROPRIATE TO BOTH CONCRETE AND FORMAL OPERATIONS ARE APPENDED. (JK)

ED016282

BRL-1646
PA24

Occasional Report No. 5

COGNITIVE STRUCTURES AND EDUCATIONAL EVALUATION

R.W. Skager and L.A. Broadbent

Center FOR THE
Study of
Evaluation
OF INSTRUCTIONAL
PROGRAMS

University of California, Los Angeles, April 1967

EA 001 020

CO-DIRECTORS

Merlin C. Wittrock   Erick L. Lindman

ASSOCIATE DIRECTORS

Marvin C. Alkin   Frank Massey, Jr.   C. Robert Pace

The CENTER FOR THE STUDY OF EVALUATION OF INSTRUCTIONAL PROGRAMS is engaged in research that will yield new ideas and new tools capable of analyzing and evaluating instruction. Staff members are creating new ways to evaluate content of curricula, methods of teaching and the multiple effects of both on students. The CENTER is unique because of its access to Southern California's elementary, secondary and higher schools of diverse socio-economic levels and cultural backgrounds. Three major aspects of the program are

Instructional Variables - Research in this area will be concerned with identifying and evaluating the effects of instructional variables, and with the development of conceptual models, learning theory and theory of instruction. The research involves the experimental study of the effects of differences in instruction as they may interact with individual differences among students.

Contextual Variables - Research in this area will be concerned with measuring and evaluating differences in community and school environments and the interactions of both with instructional programs. It will also involve evaluating variations in student and teacher characteristics and administrative organization.

Criterion Measures - Research in this field is concerned with creating a new conceptualization of evaluation of instruction and in developing new instruments to evaluate knowledge acquired in school by measuring observable changes in cognitive, affective and physiological behavior. It will also involve evaluating the cost-effectiveness of instructional programs.

COGNITIVE STRUCTURES AND EDUCATIONAL EVALUATION

R. W. Skager and L. A. Broadbent

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

CSEIP Occasional Report No. 5, April 1967
University of California, Los Angeles

The research and development reported herein was performed pursuant to a contract with the United States Department of Health, Education, and Welfare, Office of Education under the provisions of the Cooperative Research Program.

COGNITIVE STRUCTURES AND EDUCATIONAL EVALUATION

R.W. Skager
Lee A. Broadbent

It is obvious that the instructional goals around which modernized educational programs and policies are formulated must be of central concern to those who evaluate those programs and policies. This is just as true when the role of the evaluator is to provide feedback during the development and try-out of the innovation (process evaluation) as it is when the evaluator's task is to study the relative effectiveness of the finished program (product evaluation). But is it enough to evaluate instructional programs only in terms of the stated goals of the developers?

An honest, though equivocal, answer to this question is, "Yes, so long as the developer's statement of goals is reasonably complete." For example, it would be unsound to evaluate a future supersonic transport solely on the primary criteria of being able to whisk 300 people across the country in two hours. There are doubtless hundreds of other relevant standards, including safety, economy, comfort, minimum noise, and so on. Small defects in large airplanes can kill people in spectacular numbers. Aerospace developers are, therefore, likely to be rather acutely aware of their responsibility to specify all of the criteria that an aircraft ought to meet.

A very different situation arises with respect to the context of motivation underlying evaluation in education. It is apparent that unsuccessful instructional programs and policies seldom kill students. What remains unlearned today can always be learned tomorrow. Adopting a new curriculum may promote opposition from one pressure group, while not adopting the curriculum may well generate opposition from another sector. In any case, decisions about instituting educational change are likely to be made on the basis of instructional philosophy and professional salesmanship rather than in light of what the researcher would see as "hard" evidence. These and other conditions are not conducive to the vigorous analysis and energetic study of all of the dimensions involved in evaluation projects.

Another limiting factor in evaluation research may be that the developers of instructional programs often draw up rather specific instructions as to how their products are to be evaluated. Developers are naturally enthusiastic about their own ideas and sometimes rather scornful about what has been done in the past. It is especially significant that developers may reject not only earlier methods, but also earlier goals. When the evaluator is in the status of an employee or has other legitimate motivation for working closely with the developer, it is possible that the latter's enthusiasm will severely limit the thrust and scope of the questions posed in the evaluation.

These may be some of the reasons why Scriven (1965), in an important paper, has noted a widespread tendency to limit the domain of evaluation to those criteria the developer chooses to call relevant. Scriven deplores such relativism in the evaluation function on the grounds that it renders the decision as to whether a product is adequate or inadequate very difficult indeed. It is often impossible to define grounds for the objective comparison of two methods. A test designed to measure the effect of one technique is decried as being biased against the other. Yet, one cannot help but wonder why it is not perfectly appropriate to check the effectiveness of one program against the goals of another. Information relevant to more than one value system can be of genuine use to those who will eventually decide whether or not the program is to be adopted. In any case, as soon as we step back and consider the problem of instruction in a general way, the very idea that educational goals will be entirely different from curriculum to curriculum and technique to technique simply sounds absurd. It is hard to believe, for example, that the "new" and the traditional mathematics do not have at least some ends in common.

Scriven has not been the only author to suggest that the evaluator is responsible for expanding the criteria of evaluation. Though maintaining that objective comparisons between curricula will usually be of little use, Cronbach

(1963) has argued that evaluation research should cover as large a spectrum of learner characteristics as possible. We would add that it is primarily the job of the evaluator to insure such breadth, since the psychology of the developer may often center on the realization of a limited set of objectives. While some of the evaluation data may be of little immediate interest to the developer, it may be highly meaningful to users, and may---where unanticipated effects occur---be important to the developer himself.

Published Tests and the Evaluation Problem

Even though the evaluation specialist may sense that not enough criteria have been specified, he may not be certain which additional questions ought to be asked. If he does undertake a comprehensive analysis of the problem he may quickly find that tests or other measuring devices appropriate to many of his criteria do not exist. In the effort to get something done, his selection of criterion instruments will be largely determined by what is available. In effect, the state of the measurement technology establishes the criteria to be investigated instead of the evaluator.

The nationally standardized achievement test is a case in point. Its use is encouraged by the laudable contemporary emphasis on behavioral criteria. Thus, many well-meaning educators would maintain that the acceptability of a new program or technique must be defined in terms of learner achievement. In many cases it is taken almost for granted that "achievement" is to be defined in terms of performance on published tests. After all, subscale descriptions usually imply that scales measure important outcomes of instruction. National norms provide seemingly meaningful standards of comparison. Manuals indicate that panels of curriculum experts developed the test plan and inspected the items, and that great technical expertise has gone into the assembly of the final instrument. These claims are generally quite true. Nevertheless, total scores on achievement tests are less useful to the evaluator than is widely believed. Used inappropriately, the results of achievement testing can be quite misleading.

The most widely held illusion about achievement tests is that scores on such instruments actually measure achievement---that is, there appears to be a pervasive belief that achievement tests belong in a different category than do intelligence tests---that achievement tests are sensitive to what goes on in the classroom while intelligence tests are not. Examination of technical reports by reputable test publishers should rid us of this notion quickly. There are generally very high correlations between achievement tests and aptitude tests, although even the best publishers rarely comment extensively on this point.

Husek (1966) and others have pointed out that contemporary item selection procedures lead to the construction of achievement tests having psychometric characteristics very similar to tests of intelligence. Total scores are based on items selected for stability (e.g., reliability over time), meaning that achievements tests cannot be highly sensitive to the particular experiences of individual students, including those experiences provided by a given method of instruction.

Of course the point can be carried too far. If the evaluator takes time to look at the tests, he may discover subgroups of items whose content is directly relevant to the instructional programs under study. For this reason, Cronbach (1963) has urged that curriculum evaluators concentrate on individual test items rather than on total scores, especially since the latter are the summation of many skills not necessarily covered in a given program.

The criticism of published achievement tests most frequently heard attacks their putative bias toward "traditional" curriculum goals. This criticism is of lesser significance. We have already suggested that there is nothing wrong with the use of criterion measures that the developer sees as biased, although "fair" criteria should also be applied.

It is probably implicit in all views of the educational process that students not only learn skills in a given content area, but also, in some mysterious way, learn how to think. The development of a measure of critical thinking for Dressel and Mayhew's (1954) study of general education at the college level is one example of how such expectations about generalized learning influence the development of evaluation instruments. But to possess adequate generality, a description of the manner in which people think and solve problems should refer to a theoretical model. Herein lies what may be the most significant criticism of the typical achievement test. Such instruments do not give us a picture of human performance that can be referred to functions or stages in a theory of cognitive development. This is because achievement tests are usually developed by curriculum experts rather than by behavioral scientists. Perhaps, in general, this is proper. But there would be at least two significant advantages in using measures derived from developmental theory. First, as criteria, the measures would have great generality. The skills measured ought to be consistent with the long term goals of many different approaches to instruction. Second, theoretically based measures would be less likely to be perceived as biased toward a particular ideology about curriculum content. The focus of this paper, then, is to explore the literature relating to one theory of cognitive development in the hope that applications to educational evaluation will be discovered.

The Approach of Piaget

One promising theory of cognitive development is that of Jean Piaget. There is little point in describing the theory here other than in the most general terms. Briefly, Piaget's approach represents a distillation and formalization of many years of naturalistic and laboratory observation of human beings from infancy to adulthood. In contrast to American learning theory, its fundamental principles are descriptive in nature. Piaget did not begin with a set of assumptions about how learning was fixed and then proceed to generate hypotheses for experimental verification. The theory is not logico-deductive in nature, in spite of the fact that in recent years Piaget has stated it in terms of the formal model of symbolic logic. Even the principles with which Piaget began---that all organisms adapt to the environment through the processes of assimilation and accommodation---are so essentially descriptive in nature that the idea of testing their validity has little meaning. As a result, research on Piaget's theory has been concerned with verifying the observations from which the theory was derived, rather than with testing the theory itself, at least in the sense that many psychologists usually conceive of testing a theory.

Piaget has placed special emphasis on the study of the structure of intelligence. In Flavell's (1963) "structures" are defined as, "...the organizational properties of intelligence, organizations created through functioning and inferable from the behavioral contents that determine." (p. 17). For example, a child concludes that a row of seven widely spaced poker chips contains more elements than a row of seven identical chips placed close together. This judgment represents behavioral content. The content, in turn, is mediated by a cognitive structure. Piaget's task as a theoretician was to infer the organizing principles (structure) the child must have utilized in producing the particular content. These principles, accounting in this case for failure to "conserve" number, should apply to many possible behavioral contents.

Piaget concluded that certain sets of organizing principles operated at certain ages, to be replaced later on. Hence, he organized his theory according to a series of developmental stages. The theory, then, can be viewed as an increasingly complex hierarchy of cognitive structures describing how it is that human beings produce particular behavioral contents at particular ages. Cognitive development, in the Piagetian sense, is not analogous to the slow accretion of knowledge, as might be the model for the construction of achievement tests. Rather, the theory refers to progression through a series of levels with qualitative differences existing between the structures employed at each level. The appearance of new behavioral content requires the postulation of new, or at least elaborated, principles of dealing with the environment.

Paradoxically, Piaget's work has not been selected because the theory is at its strongest when explaining how particular cognitive operations are learned or acquired. The theory deals with this issue in general terms, but suffers from a lack of particularization with respect to the mechanism by which specific cognitive structures actually develop. Instead, Piaget's approach has been selected because of the wide variety of cognitive operations it incorporates. Equally important, Piaget has devised literally hundreds of tasks eliciting the behavioral contents his theory was developed to explain. It may be that in devising these tasks, e.g., in his choice of what it was that would be observed, Piaget made his greatest contribution to developmental psychology. The tasks, and the questions raised by the resulting behavioral contents, frequently have a way of convincing the reader that he is looking in on something of genuine importance. Flavell (1963) has suggested that these tasks amount to "...literally a warehouse full of interesting and face-valid measures of intellectual performance..." (p. 417). Thus, it may be possible to develop a battery of measuring instruments derived from a theory of cognitive development. We are not particularly concerned here with the validity of Piaget's theory or even with the correctness of some of his observations. For example, for present purposes, it is not important whether conservation of space invariably precedes conservation of volume.* Both concepts refer to aspects of cognitive functioning that seem to be important milestones in the development of intellectual functions. Thus, Piaget's observations should be of interest even to those who do not admire his theory.

In the sections which follow, we will be concerned with what the literature has to say (or, more often, what it fails to say) about three questions that need to be answered before the feasibility of applying Piaget's tasks to the evaluation of educational programs can be established. The three questions that must be explored are:

1) What are the tasks themselves like, and which ones are of particular interest?

---

* Much of the research on the theory is concerned with such issues. In one of the more extensive studies, Dodwell (1960) reported that 5-8 year old children showed little within subject consistency in level of conceptual development across five different conservation tasks. This finding is in disagreement with those reported by Piaget.

2) Is there evidence that performance on any of the tasks can be influenced by educational or cultural factors?
3) Is there any evidence that the tasks measure skills that are sufficiently independent to warrant their further development? That is, Piaget's tasks look different, but are they really?

The Tasks

Piaget and his co-workers have always collected their information about cognitive functioning through unstandardized, face to face encounters with individual subjects. Except during the period of infancy, these encounters generally follow the format of confronting the subject with some sort of problem situation. The tasks are unstandardized in that individuals are treated differently depending on their initial responses. Great efforts are made to get the maximum level of performance out of each subject. If a right answer is given, the experimenter always asks why or how.

Because the focus of this paper is on the application of Piaget's ideas to educational evaluation, we are not particularly interested in measuring aspects of cognitive structure in infants or pre-school children. We are concerned with two broad categories: the stage of concrete operations, extending from approximately seven to eleven years, and the stage of formal operations, beginning between age twelve and fifteen and culminating in adulthood.

"Concrete operations" defines, in part, a period in which the child develops certain important representations of the physical world, such as the quantitative concepts of matter, weight, volume and number. It is during this stage that the child's judgmental processes become increasingly independent of irrelevancies introduced over dependence on direct perceptual cues. This independence is achieved through the formation of differentiated quantitative concepts, illustrated by performance on the conservation tasks (certainly the best known of the Piagetian procedures). Conservation of matter is attained, for example, when the child recognizes that the amount of a substance is not influenced by its external shape. Conservation of number is indicated when the child recognizes that the way in which a set of objects are distributed in space (e.g., spread out or bunched together) is unrelated to how many objects there are.

A second characteristic of children at the stage of concrete operations is a growing ability to utilize two types of logic: the operations of class inclusion and serial ordering. Class inclusion operations are formalized in the familiar model of arithmetic, where, for example, part-whole relationships are manipulated by addition. In Inhelder and Piaget's (1958) example, "boys plus girls equals children" amount to a class-inclusion operation. Serial ordering operations can be represented, in part, by the algebra of inequalities. The ability to arrange objects in a series according to some attribute and to find similarities between series are generalized examples of these types of structures. Incidentally, "operations" have been defined by Flavell (1963) as representational acts which are integral parts of an organized network of related acts. The actions implied in the common symbols of arithmetic are all examples of generalized intellectual operations applicable to relationships that do not involve numbers.

Most investigations at the level of concrete operations have concentrated on relatively few of Piaget's tasks. Conservation of matter, weight, volume, and number have been studied by means of procedures that are becoming relatively familiar. Conservation of matter, for example, is usually determined by changing the shape of balls of plasticine or clay and asking the child whether or not the modified object contains as much material as the original, as did Lovell and Ogilvie (1960), Smedslund (1961a, 1961b, 1961c, and 1961d), and Pratoomraj and Johnson (1966). Conservation of volume, also studied frequently, is measured by a variety of tasks such as pouring liquids into vessels of different shapes, as illustrated in Bruner (1964), or by rearranging blocks built from smaller elements, as did Lovell and Ogilvie (1961). Conservation is indicated when the child recognizes that size and shape of the container do not influence the amount of space occupied by the substance.

Conservation of volume is the last of the quantity conceptions to be attained according to Piaget, and this has been verified, for example, by Uzgiris (1964). In spite of the fact that the "normal" child should have attained conservation of volume by about the age of twelve or even earlier, Elkind (1962), in a study of American college students found that only a little more than half of the subjects, all of whom should have been at the stage of formal operations, had attained this quantity concept. This finding is so inconsistent with Uzgiris (1964), who observed the mean age of attainment to be 10.7 years, as well as with Lovell and Ogilvie (1961), that it is difficult to accept. However, it may be that there are enough individual differences in age of attainment of various quantity conceptions to warrant their use for a far wider age range than might be expected.

Relatively few investigations have been concerned with the logical operations attained during the stage of concrete operations. Smedslund (1960), in a study of

5-7 year olds, established simple preference orderings for pairs of objects. Tests were then made for transitivity to two objects that had not been paired previously. For example, if A is better than B, and B is better than C, can the child derive the ordering for the pair A and C? Smedslund's results were negative with respect to transitivity of serial ordering at all the ages studied, as is consistent with Piaget's observations. These findings were elaborated by Smedslund (1963) in a later study.

Appendix A provides a summary of the tasks encountered in our survey of research at the level of concrete operations. While these do not cover all possible tasks, they are certainly the ones used most frequently.

Cognitive functioning at the stage of concrete operations is characterized by one very significant limitation in that the operations utilized in solving problems are mainly confined to dealing with the actual or phenomenal rather than with the possible or hypothetical. Distinctions between cause and effect are made primarily on the basis of the summation of empirical observation rather than via the systematic formulation and testing of hypotheses. Inhelder and Piaget (1958) dwell at some length on the performance of children of different ages on a kind of billiard game that illustrates rather well the transition period between concrete and formal operations. (This apparatus is described in Appendix B as the "Equality of Angles of Incidence and Reflection Problem"). A plunger that can be swiveled in an arc projects a ball against a barrier. The subject shoots the ball off the barrier in an attempt to hit one of a set of targets moved about from trial to trial. The task is to arrive at a rule by which the ball can be made to hit the target. Younger children up to the age of 7 or 8 years are unable to arrive at formal rules for success, even though performance may improve with practice. The situation is perceived in a global and undifferentiated fashion. For example the ball is viewed as moving in an arc rather than in the two separate segments of incidence and reflection.

Somewhat older children at the stage of concrete operations are able to formulate essentially empirical rules, involving statements such as, "To get over there you have to aim over here." Although these children apparently see the ball moving through an angle rather than an arc, the angle is not broken into segments. In contrast, the formal reasoning adolescent takes a hypothetical approach to the problem. It is reported that eventually all subjects at the formal level discover that by aiming the plunger directly at the barrier, the ball can be made to rebound to the point from which it was originally projected. This observation is particularly useful in formulating a rule involving equality of angles of incidence and reflection. Given the rule, the formal reasoner is able to test its validity in one or two additional trials.

Appendix B contains brief descriptions of the formal reasoning tasks presented in Inhelder and Piaget (1958). Although these tasks vary in complexity, all require the subject to do at least two things: (a) differentiate the variables in the problem, and (b) assume a rule-generating, hypothesis-testing approach independent of empirically oriented manipulation of the apparatus---that is, given familiarity with the elements of the apparatus, the subject has to be able to isolate critical manipulations at a symbolic level. This ability to rise above the phenomenal defines the essential difference between the stages of concrete and formal operations. The older subject can perform a "combinatorial analysis"---an important concept in Piagetian theory---or putting-together of the various potential casual relationships involved in a problem.

The survey of the literature relating to Piaget has revealed that there has been very little empirical work outside of Geneva at the level of formal operations. In part, this may be due to the fact that the complete elaboration of the theory at the formal stage is so recent. It also may be that the propositions in symbolic logic that characterize the formal theoretical model are likely to have frightened off many an aspiring investigator. In contrast, the literature concerned with verification of Piaget's observations at the level of concrete operations is by now quite extensive. This lack of interest at the higher level is more apparent than real, however. Almost all of the vast literature on problem-solving and much of the literature on concept formation deal with performance on tasks that can be assigned to the level of formal operations. Most investigators at this level speak theoretical language different from Piaget's. An interesting critique of some Piagetian principles has been provided recently by Berlyne (1965), who presents a model of his own integrating much of Piaget's theory and observation.

Studies of the Effects of Training on Cognitive Development

The environmental orientation of many developmental researchers is doubtless at the root of the considerable interest in the effects various training procedures have on performance on some of Piaget's tasks. There is a general impression that Piaget's own orientation is contrary to the environmentalist position. This is probably valid, even though one would be hard put to find a definitive statement in the matter in any of his writings. That culture and education effect cognitive development is certainly admitted, though this topic is evidently of minor interest to the Genevans.

All operations develop through "actions" with the environment, but endogenous factors might well place rather restrictive limits on when and how such structures develop. Piaget's actual position on the matter can be inferred from the Inhelder and Piaget (1958) book. In describing the performance of subjects on the Floating Bodies Problem (Appendix B), these authors suggest that "...it is hard not to allow for the role of spontaniety in the progressive structuring of the data, even if it is hastened by the surrounding social environment." (p. 37).

The evidence for cultural and training ~~of~~ effects is obviously critical if any of the Piagetian concepts are to be useful in educational evaluation. If performance on tasks eliciting the various quantitative concepts and logical operations cannot be influenced by laboratory training designed specifically for its facilitation, it would seem unreasonable to expect that more generalized educational policies and techniques could exert any influence. Cross-cultural comparisons for differences in level of cognitive development are at least of equal importance. Experimental subjects spend very little time in the laboratory as compared to the magnitude of the influences exerted by the experiences of their everyday lives. Perhaps many educational variables are more in the cultural than the laboratory category. There has, nevertheless, been very little cross-cultural work on Piagetian concepts, in spite of the fact that many writers seem to believe the matter to be settled already. In their recent review, Ausubel and Ausubel (1966) conclude that cultural and experiential factors have been demonstrated to have greater influence on Piagetian concepts than do age and intelligence level. Unfortunately, a careful reading of the sources cited by the Ausubels in support of their assertion fails to show anything of the kind.

Much of the empirical work on training effects has contrasted the effects of what is usually termed "reinforcement" against procedures designed to provide understanding of principles. For example, Wohlwill and Lowe (1962), studying conservation of number in 5 and 6 year old children, had Ss under the reinforcement condition count the number of elements in a set in order to learn that quantity is unaffected by spatial rearrangement. Another group was given the same type of practice along with experience in the additional and subtraction of elements, presumably demonstrating that these operations are the only means by which the number of elements can be changed. A third group was given training designed to show directly that alteration in the configuration of the set of elements did not change their number. The results of this carefully controlled experiment were disappointing. There was no evidence for transfer from any of the training procedures to a measure of conservation of number. The learning that did occur was apparently restricted to the training sessions themselves.

The most extensive series of studies along these lines has been conducted by Smedslund, who investigated conservation of weight in an initial study (Smedslund, 1961b) similar to the Wohlwill and Lowe work. Under the reinforcement condition, children 5 to 7 years of age were trained by weighing objects on a balance after the form of one had been altered. Another group was given a training procedure in which the form of the objects remained constant while weight was changed by adding or subtracting small bits of material. Again, neither experimental conditions appeared to induce conservation responses. Smedslund (1961c) later was able to train children to give responses indicating conservation of weight by means of the same reinforcement procedures described above. The fact that this achievement was incidental to the purpose of the later study does not explain the inconsistency with previous work. Smedslund demonstrated that a group of children who had acquired conservation of weight under the reinforcement condition quickly lost the concept under extinction trials, while about half of a group of children who had already achieved conservation on pre-test refused to give up the conservation concept in the face of conflicting evidence artificially contrived by the experimenter. In this case "extinction" involved the covert removal or addition of small pieces of material at the same time that the shape of one object was changed. Smedslund suggested that the reinforcement training merely taught the subjects to give correct responses in a particular situation, without really achieving the conservation concept. Once conservation is truly learned, subjects refuse to accept negative evidence, searching for other explanations such as, "A piece must have dropped to the floor." In a related study, Smedslund (1961d) also included an extinction condition, but this time in the attempt to teach rather than extinguish the conservation concept. Six-year-old subjects were given 36 trials at weighing pairs of objects that differed in size but not weight. This procedure was designed to reduce the children's dependence on irrelevant cue of size. Again, the training condition was ineffective in inducing conservation.

The last two studies in the series did provide positive evidence that training could increase conservation responses. In a complex study, Smedslund (1961f), confronted subjects with a choice situation eliciting conflicting responses. For example, at pre-testing, a given child believed incorrectly (a) that elongating an object increased the amount of the material it contained and correctly (b) that removing part of the object decreased the amount of material during training. The experimenter would perform both of these operations on the same object and then ask the subject whether the amount of material in the object had changed. Smedslund's hypothesis, derived from Piagetian theory, predicted that the conflicting responses induced by this choice would create a state of disequalibrium in which the child would be more likely to attend the operation of addition and subtraction, disregarding the irrelevant cue of shape deformation. Only 13 subjects participated in the experiment, all non-conservers on pre-test. Five subjects consistently attended the addition-subtraction cue during training. Of these, four attained conservation of the post-test. Of the eight subjects who chose the deformation cue during training, none attained conservation. Smedslund considered this finding to be the first convincing evidence that conservation can be trained. In a second experiment Smedslund (1961f), using a control group, applied the same training procedures to conservation of substance with both continuous material (plasticine) and discontinuous objects (colored chips). This is one of the few training studies in which the experimental group exceeded a central group on a conservation post-test.

Bruner (1964) describes an attempt by Frank at training conservation of volume, using the usual technique of pouring liquids into containers of various shapes (see Appendix A). Children aged 4, 5, 6, and 7 were pretested on conservation of volume. Age-wise, the subjects varied in the proportion initially giving conservation responses from none of the four year olds to approximately half of the 6 and 7 year olds. The training procedure involved concealing the differently shaped containers while pouring in the liquid, thus removing the irrelevant perceptual cue. Under this condition, all of the 6 and 7 year olds gave conservation responses, as did 80 per cent of the 5 year olds and 50 per cent of the 4 year olds. When the screen was removed, the 4 year olds returned to the non-conservation responses, while the older children gave considerably more conservation responses than occurred on the pre-test. The post-test was given only a few minutes after training, so a question remains as to how enduring the training effects may be. It does appear that the rather simple expedient of removing the irrelevant cue of shape increases the frequency of conservation responses.

Inhelder, Bovet, Sinclair, and Smock (1966) have recently disputed the above findings. These investigators confronted children with the discrepancy between a non-conservation response and the actual facts. This was accomplished by means of an apparatus demonstrating that fluids drained out of two jars of different shapes occupied the same volume in two identically shaped jars placed underneath. A second procedure was designed to promote the generalization of an existing structure to conservation of volume. The subjects were already equipped with the "conservation of discrete unities" concept (altering the configuration of a set of beads does not change their number). During training, these subjects observed sets of beads being poured into containers of different shapes. The beads became successively smaller, eventually approximating the action and consistency of a fluid.

The authors reported that some subjects made more conservation responses on post-test, but maintained that these were individuals who initially gave "subtle" signs of being on the verge of attaining conservation of volume. There was little or no evidence that the training persisted on a post-test given two weeks later, when countersuggestion by the experimenter eliminated any tendencies to give conservation responses. In view of these negative findings, Inhelder, et. al., disputed the validity of Frank's research. While the latter most assuredly should have included a more delayed post-test, the two studies were so different that it is simply gratuitous to dispute one on the basis of the other.

The majority of the training procedures above appear to require that subjects have some understanding of the operations of measurement, whether implicitly as when comparisons are made between occupied volumes, or explicitly as when different objects are weighed on balance scales. If the relevant measurement operations are unavailable to subjects, it is hard to see how conservation training can be effective. In an initial study, Beilin and Franklin (1962) were successful in training first and third grade children in the measurement of length, but found that area measurement could be learned only by the older group. It is unfortunate that these authors did not extend their research by relating conservation of length or area to the availability of related measurement operations.

In a later study, Beilin (1965) administered various conditions (one of them derived from Piagetian theory) in the attempt to train conservation of length and number. A test was then made to determine whether learning transferred to conservation of area. "Convergence" was defined as an increase between pre-test and post-test in the correlations among the three types of conservation tasks. The only clearly effective training procedure turned out to be direct instruction in verbal rules or conservation. Under this condition, the experimenter simply repeated the conservation rule each time the subject failed to give a conservation response. While there was an increase in conservation responses for length and number, there was no evidence of transfer of training to the conservation of area. Evidence for convergence was confined to length and number. The correlation between these two tasks rose from .44 to .72 on a post-test administered after a period of several weeks.

In view of the ingenious and often complex training procedures adopted in other studies, it is interesting that Beilin's research is the only instance where one of the conditions involved the simple and doubtless naive expedient of telling the subject that changes in configuration, etc., have no effect on quantity. The efficacy of this procedure is of theoretical importance. Bruner (1964) argued for a relationship between the development of the conservation concepts and the child's attainment of the ability to represent past events on a symbolic level. Inhelder, Bovet, Sinclair, and Smock (1966) took strong exception to the suggestion, denying that language learning *per se* is involved in the organization of informational units leading to conservation responses. Piagetian theory emphasizes the "interiorization of actions" rather than the mediating role of representational models. Beilin's findings with regard to the effectiveness of instruction in verbal rules obviously supports Bruner's position.

Evidence for Cultural Effects

As we have indicated, there is relatively little on which to base any conclusions as to the role of cultural factors in conceptional development. In a study involving a relatively small number of subjects, Dodwell (1961) reported that socio-economic status and rural vs. urban origin were related to performance on an objective test designed to measure Piagetian number concepts. Smedslund (1961a) reported that Geneva children from a superior socio-economic level were found to achieve conservation and transitivity operations somewhat earlier than did children in general. Smedslund also mentioned two unpublished English studies supporting the notion that environment influences the development of number concepts.

Goodnow (1962 and 1966) conducted important cross-cultural studies comparing Chinese children and adults of various educational and socio-economic levels with a group of European children. Three conservation tasks were used in the first study, (space, weight, and volume) along with an additional task requiring formal operations. The latter required the subjects to make systematic combinations among a set of chips of different colors. While the earlier study bore on many issues other than cultural effects, one finding appeared to reflect the effects of experience on conceptual level. Among the Chinese, a group of students taking highly theoretical, non-laboratory curricula in science did more poorly on the conservation tasks than students in more practically oriented courses. The investigator suggested that this was because the former students, perpetually confused by their educational experience, had learned to distrust their own judgments. In general, however, Goodnow interpreted the results of the first study as showing little evidence for milieu effects.

In the second study, Goodnow (1966) compared one of the Chinese groups with a number of American groups differentiated according to mental and chronological age. The finding that concerns us particularly here is that, while the unschooled Chinese did about as well as American children of similar chronological and mental age on the conservation tasks, they performed far more poorly on the combinations task. Goodnow suggested that the formal requirements of the latter task---that "things be worked out in the head"---give the advantage to the child who has been to school. In contrast, the experiences that contribute to the learning of conservation may be so common within any human cultural environment that little can be expected in the way of differential effects.

Goodnow's findings are particularly important in light of the many studies on training effects cited in the previous section. All of these studies were conducted at the level of concrete operations, and were further confined in the main to the various conservation concepts. The stage of formal operations, where there may be more optimistic grounds for expecting educational and cultural effects, has been almost completely neglected.

Our review of the literature on training effects and cultural differences can be summarized in two statements.

(1) Evidence for the effects of training on cognitive development at the level of concrete operations is,

at best, equivocal. The majority of investigators, in fact, report negative results. However, particular training procedures may be effective, such as those reported by Bruner, (1964), Beilin (1965), and Smedslund (1961e).*

(2) Conceptual development at the level of formal operations has thus far been virtually ignored. This gap in our knowledge is to be deplored, since there is some reason to believe that educational and cultural experiences have greater influence at more complex levels of performance. The level of formal operations appears to be a suitable field for future investigations into factors relating to the development of the cognitive operations comprising Piaget's system.

Subject Characteristics

An overview of empirical studies conducted to date on Piaget's theory suggests rather forcibly that most of the investigators have had little interest in how performance on Piagetian tasks might be related to subject characteristics other than chronological age. The paucity of cross-cultural studies is one testimonial to this observation. This situation is an unwholesome one, because the failure to relate subject characteristics to level of conceptual development has the obvious disadvantage of not allowing for the discovery of relationships that will aid in the integration of Piaget's findings with observations made under other theoretical orientations.

---

* The Educational Testing Service (Bussis, 1965) is presently experimenting with use at the primary level of materials designed to teach logical operations incorporated in Piaget's theory. For instance, children are given exercises in the temporal ordering of series of pictures, presumably requiring the serial ordering operation. This project is especially interesting because the materials have been designed for classroom rather than laboratory use.

Goodnow's (1962) comparison of Chinese and Europeans provides an example. The Chinese subjects evidently applied more precise standards in their equality judgments, resulting in differerences between groups from the two cultures on several measures. This finding suggests that the conceptual style of equivalence range, described by Clayton and Jackson (1961) as a tendency to use broad and inclusive vs. narrow and exclusive conceptual categories, may be related to performance on conservation tasks.

The usual conservation problem involves a judgment as to whether two quantities are equal, e.g., whether assignment is to be made to the same or different categories. The subject who customarily uses narrow categories would presumably be more likely to make a judgment of non-equality, thus performing poorly on conservation problems. We also learn in the Goodnow study that adult subjects who performed less accurately on conservation tasks scored lower on the block design test of the Wechsler-Bellevue. Witkin, Dyk, Faterson, Goodenough, and Karp (1962) reported high relationships between this subtest and the personality characteristic of field independence. On the basis of Witkin's conceptualization we would indeed expect field-dependent people to pay more attention to external as compared to internalized cues. In the conservation tasks the irrelevant perceptual cues may be misleading in a manner that is at least analogous to some of the situations used by Witkin and his collaborators. Thus, we would expect adults having trouble with conservation concepts to be more field-dependent.

For present purposes, we are especially interested in relationships between aptitude measures and performance on the Piaget tasks. If such relationships are high, there is reason to doubt that the tasks will be sensitive to the influence of instructional variables. While aptitude tests admittedly measure nothing more than what the subject has learned, the stability of measured aptitude over time makes such tests poor candidates for indicating the effects of particular learning experiences.

Relationships between aptitude measures and level of conceptual development were reported in a few of the studies reviewed. Dodwell (1961) administered individual tests of number concepts modeled after the Piagetian approach and found that scores obtained during kindergarten correlated .59 with performance on an arithmetic achievement test given at the first grade level. We can probably assume that the achievement measures, in turn, would be highly related to a measure of general intelligence.

Case and Collinson (1962) designed an interview situation to elicit responses at the intuitive, concrete, and formal leads of conceptual development. Subjects aged 11 to 18 responded to questions about reading passages that referred to conflicting aspects of the material read. A scoring scheme to be described later allowed for assignment of responses to one of the three conceptual levels. The scores derived from this technique correlated .65 with mental age in the Raven Progressive Matrices and .66 with chronological age. These correlations are quite high, though the task admittedly differs considerably from the measures of formal operations used by Piaget.

Goodnow (1962) originally reported that means on the conservation tasks did not appear to be related to group differences in aptitude. On the combinations task, requiring formal reasoning, the status of the groups did parallel scores on the Progressive Matrices Test rather closely. In the later study, Goodnow (1966) reversed the first of the findings by reporting the performance of groups of American children on the conservation tasks was highly related to mental age. Bright 8 year olds, presumably having less experience, performed better than did dull 11 year olds of lower mental age. Goodnow did not report correlation coefficients.

Freyberg (1966) also reports the development of a test of Piagetian number concepts. His measure was found to correlate .52 with a mental age estimate from the _Primary Mental Abilities Test_ and added only a tiny increment to the PMA in the prediction of later performance of tests of achievement in arithmetic. Again, the "concept" test was a different sort of measure than those used by Piaget, in spite of the theoretical basis on which the test was developed.

Summary

There is very little evidence as to the relationships between performance on Piaget's tasks and standardized measures of general intellectual capacity. What evidence there is covers a very limited set of tasks---again mainly ignoring the level of formal operations---and has been provided in the main on samples of subjects selected for special purposes. While the issue thus must remain open, particularly at the level of formal operations, there are at least preliminary grounds for anticipating that correlations between aptitude measures and performance of Piagetian tasks will prove to be substantial. This is not an encouraging report with respect to the potential utility of the tasks in the evaluation of instructional programs.

Differentiation in Cognition

Little has been said in this paper about the practicality of most of Piaget's tasks. The use of apparatus and time-consuming individual administration as well as the need for experienced administrators, hardly suggests that the use of the techniques on a mass basis will be feasible. In reality, this problem is not a matter for serious concern at present. The first need is to establish whether the original measures are meaningful and useful. Psychometricians have often been able to develop practical methods for measuring variables originally defined in complex, laboratory situations. For example, Jackson, Messick, and Myers (1964) reported a group embedded figures test to be highly related to an individual test originally used in defining the field-independence personality dimension.

There is much more to the matter of correlates of the Piaget tasks than mere practicality of administration, however. Descriptions of the way in which subjects deal with various problem situations often give the impression that subjects who fail are simply unable to isolate the various elements of the problem. In other words, subjects who perform badly are not analytical and differentiated in their cognitions. There are many examples. In Beilin's (1965) study of the conservation of length the experimenter constructed straight-line figures by placing sets of small sticks of equal length end to end. The non-conserving child had no difficulty discriminating which figure is the longer when the sticks were so arranged. Yet when one of the figures was turned into an "L" by putting the last stick at right angles, the non-conserving child saw it as the same length as a straight line figure with one less component stick. It is apparent that such a child perceives the bent stick as a gestalt rather than as a total length differentiated into two additive elements. In the sense in which he perceives the problem, the non-conserver is quite correct. The bent stick qua figure is the same length as the shorter of the two straight sticks.

Case and Collinson's (1962) study provides another example. It will be recalled that the subjects responded to questions on written material that provided conflicting or equivocal information. To be scored at the formal level, responses had to be analytical and differentiated, taking inconsistent aspects of the material into account.

Similarly, though at a lower level, Piaget's much-used principle of compensation implies that the subject able to conserve volume take into account that that liquid poured from a short, wide container into a tall, narrow container increases in height while decreasing in diameter. This and other conservation responses apparently require a differentiation of the elements of the stimulus situation.

The pin-ball game (appendix B) illustrating the principle of equality of angles of incidence and reflexion provides a final example. It was pointed out earlier that the equality principle is generated when the subjects are able to break the arc of the ball into two discrete segments. Inhelder and Piaget (1958) emphasized that younger children did not perceive that the path of the ball followed an angle, but rather described the motion in terms of a curve. In effect, the appropriate elements of the situation were not differentiated.

In the debate with Bruner about the primacy of language in cognitive development, Inhelder, et. al., (1966) reported on an analysis of the verbal output of conserving and non-conserving children. It is quickly apparent in the discussion that conserving subjects used language of a precisely differentiated nature. For example, conservers frequently utilized relational expressions such as "more than," rather than describing each stimulus uniquely. Conservers used "couples of opposites" such as "long/short" and "fat/thin," rather than overlapping couples such as "big/small" and "large/small." Moreover, conservers more frequently applied "coordinated descriptives," terms that referred to multi-dimensional rather than uni-dimensional stimulus characteristics. These observations indicate that conservers use more comparative, differentiated, and analytical language.

Tests that require subjects to describe, contrast or categorize stimuli with multidimensional attributes thus may be closely related to performance on many of Piaget's problem situations. Such tests would be scored in terms of the number of stimulus attributes noted by the subject, or for the type of language used in description, or for the differentiation of a conceptual system produced by the subject. If measures of this type are found to have substantial correlations with performance on Piaget's original tasks, it may be possible to use them in estimating level of cognitive development in large scale research where administration of the original problem situations is not practical. It would also be of great interest to find out how such measures are related to tests of the analytic vs. global conceptual style as elaborated by Kagan, Moss, and Sigel (1963); the equivalence range style studied by Clayton and Jackson (1961) and Messick and Kogan (1963); and the category width style discussed by Pettigrew (1958). Such research should also contribute something to the understanding of performance on the original Piagetian tasks.

Implications for Evaluation

It is clear that the usefulness of Piagetian concepts and tasks in the evaluation of instructional programs has not been established by research to date on cultural and training effects. There is at least some evidence that laboratory training influences the learning of conservations concepts, although Flavell (1963) was not optimistic after his summary of research published before the year 1962. Similarly, the lack of information as to the relationship between Piagetian measures and constructs developed under other theoretical orientations leaves the theory rather isolated, lacking that net of relationships that might tell us something about the generality and meaningfulness of the measures employed.

In the most important sense, however, it is the nature of the theory itself that generates interest in the measures. Hunt (1961) has emphasized that Piaget's conception of the order of appearance of the various cognitive structures, if corroborated, will supply the basis for the first "natural ordinal scale of intelligence." Ever the environmentalist, this author points out that, "...the existence of such a natural ordinal scale would suggest the investigative method of measuring the effects of various kinds of experience on the rate of development..." (p. 356). Thus, we are back to where we started. It is apparent that the most direct way of determining the usefulness of Piagetian concepts in educational evaluation is to use his measures in actual evaluation studies.

Piaget's theory, and hence Piaget's measures, will not be applicable to every type of instructional program. There should be a reasonable basis for believing that the program is designed to influence the development of cognitive skills. It should also be emphasized that measures of fundamental cognitive skills may be very useful in evaluation studies, even where performance on the measures is not found to be influenced by particular educational practices. For example, the acquisition of cognitive operations such as class inclusion and serial ordering might differentiate between those who are able to profit from an instructional program in mathematics and those who are not. In this sense, the measures would provide important interpretive information not obtainable from tests measuring developed skill at dealing with the number system itself.

In view of the gaps in knowledge outlined in this paper, the use of Piagetian measures in appropriately designed evaluation research may well have pay-off transcending the decision maker's immediate goals. Such evaluation studies, if conducted with the theoretical literature in mind, will certainly provide important data on age ordering, training effects, correlations with other measures, and cross-cultural differences. Thus, evaluation research, guided by essentially pragmatic goals, can be just as effective as a tool in the investigation of fundamental issues in developmental theory as can what is usually referred to as basic research.

APPENDIX A

Tasks Appropriate to Level of Concrete Operations

The tasks summarized below by no means constitute an exhaustive list of descriptions of Piaget's testing procedures at the level of concrete operations. Such a list would approach book length. Instead, these are procedures encountered in examining research relating primarily to studies of the effects of training on level of conceptual development. However, the list is fairly representative of the selective emphasis given contemporary research to certain aspects of Piaget's theory. It will be apparent that conservation tasks (quantity, number, etc.) are much more adequately represented than are studies of concrete operational structures, such as the logical operations involved in grouping, operations on classes and relations, and the related structures for arithmetic operations. The summaries below, therefore, leave out many measures of potential use in the evaluation of instructional programs.

Conservation Tasks

Although further subdivisions could be made among the tasks listed under this heading, all in one way or another involve the issue of whether certain physical properties of objects remain invariant in the face of changes or transformations in other properties. Such invariance (e.g., "conservation") is required if concrete operations are to be performed on the objects or quantities.

Length: Piaget, Inhelder, and Szeminska (1960), using a number of blocks and a table as materials, had Ss build towers equal in height to those built by E. The S was first to reproduce a tower built by E on the table. To test for conservation of length (or height), E constructed a second tower on the floor, the top of which extended above the table. Then, S was asked to build his tower on the table equal in height to E's tower. In order to achieve conservation of length, S had to take into account the higher level at which his tower was begun.

Smedslund (1963) studied conservation of length using two sticks of wood as materials. While stick A was actually longer than stick B, the two sticks were laid on V-shaped cardboard arms in such a way as to produce a Muller-Lyer Illusion with stick B appearing to be longer than A. Both sticks were raised off the arms to a vertical position (eliminating the illusion), and S was asked to judge which stick was longer. Conservation of length was demonstrated when S was able to resist the influence of the illusion.

Beilin (1965) used a different method. The materials were 1-inch sticks (¼ inch in diameter). Three parallel columns of straight lengths were laid out (approximately 8 inches apart, with the top ends aligned) on a square cork board (2x2 feet). The sticks were red and round during the pretest and training series and were green and square during the post-test. In all other respects, the pre- and post-tests were identical. The length conservation test consisted of 2 practice and 12 test trials, with each trial divided into two parts. First, S was shown three parallel columns, each constructed of sticks placed end to end. One of the outside columns was equal in length to the middle column, and the other was either longer or shorter than the middle one. The S chose the left or right column "like" the middle one by pressing a nearby button. On the second half of each trial the S watched the E change the two outer columns by aligning the beginning and terminal points of the incorrect column with the middle column. This was accomplished by laying out the "incorrect" column in the form of an "L" with the "end" aligned with the middle column. The "correct" column was made perceptually different from the stimulus column by having its terminal point _not_ coincide with the stimulus length (i.e., by similarly making it an L-shaped figure). After the change, S was again asked to choose the row "like" the middle one. Length combinations were changed in each trial. Conservation of length was demonstrated when Ss picked the column of the same absolute length as the middle column, disregarding the fact that it was perceptually shorter.

Weight: The classical test for the conservation of weight utilizes balance scales and balls of plasticine. The S is given two identically shaped balls of plasticine and informed that the balls are of equal weight. The shape of one is then modified by E. The task is to determine which of the two quantities is heavier (or lighter) and to indicate why. (This procedure was used by Smedslund, 1961a, 1961b, and Goodnow, 1962.) Ss reporting that the weight remains the same regardless of shape have achieved conservation of weight. Major variations in this task focus on using different kinds of materials and procedures. Here, substance sounds like weight. (Smedslund 1961c, 1961d).

Substance: The usual test for this quantity concept also uses balls of plasticine, though without balance scales, and S judges which object contains more or less material. The shape of the plasticine was changed to determine the stability of conservation (Smedslund, 1961). This basic procedure was followed by Smedslund with minor variations in kinds of materials. In contrast to plasticine, a "continuous" material, pieces of linoleum or other substances are referred to as "discontinuous" materials. For example, two equal piles of yellow linoleum squares are compared. One of the piles is then changed into a cross. Conservation is indicated when it is reported that the two piles contain same amount of material, regardless of perceptual differences.

Lovell and Ogilvie (1960) also used balls of plasticine in studying conservation of substance. S was asked to select two balls of exactly the same size from an assortment of balls of various sizes and colors. One ball was then rolled into the shape of a sausage and the child was tested for conservation. Ss not demonstrating conservation on the first test witnessed pieces of material being added to/or subtracted from one of the balls, followed by questions as to whether the two objects now contained the same amount of material. Later trials with the deformation procedure indicated that the addition and subtraction method helped clarify the conservation concept for at least a few of the Ss. A second experiment was conducted with a different material to test for the generality of this quantity concept. A rubber band was shown to the child. It was then stretched and S was questioned as to whether there was the same amount of rubber in the band as before.

Uzgiris (1964) used the following materials and transformations for investigating conservation of substance: (a) equal plasticine balls, one then changed to a sausage, a long cylinder, and finally 3 pieces; (b) metal nuts built into the form of two structures each containing 18 elements, with one arranged 3x2x3 and the other in the form 3x3x2; (c) a row of nuts 3 elements long and 6 high versus three piles of 3x2; (d) coils of braided wire, one unstretched and the other stretched almost straight, then a second condition where one-third the strands of the straight wire were separated to form two pieces; and (e) two pieces of straight plastic wire with one then tied into a simple knot, tied a second time and twisted to a round shape, and finally, straightened and cut into three pieces. In each case, the S was asked to tell which object contained more or less material and why.

Pratoomraj and Johnson (1966) examined the effects on conservation of substance of the form of question asked (e.g., "Is it the same?" vs. "Is it different?"). In addition to using materials such as clay and blocks, they used glasses of water, blocks of equal length first placed equidistant and then with a brief case placed between two of them.

Area: Goodnow (1962) presented the Ss with two grass "fields" with houses placed in each. There were an equal number of houses on each field, bunched together in one case and scattered apart in another. The task was to judge whether a horse has more grass to eat when houses are bunched together or when they are scattered apart. Ss showing conservation of area would indicate that the same amount of grass was exposed in each field.

Using a more elaborate procedure, Beilin (1965) altered the typical conservation of area task, referring to his procedure as a measure of "quasi-conservation." A Visual Pattern Board (VPB) apparatus consisted of a display panel and a control box. By inserting a prepatterned template into the control box, it was possible to display any pattern of lighted areas within a 12x12 matrix of squares. Three kinds of patterned pairs were presented. In one series, the areas were equal and spatially congruent; in the second series, the areas were unequal and spatially noncongruent (e.g., 4 squares vs. 5); the third (quasi-conservation) series consisted of pattern pairs, equal in area, with transformed configurations which made them spatially noncongruent. The S was presented with a pair of light patterns on the VPB and asked whether the two figures covered the "same" or "different" amounts of space. In the area quasi-conservation trials, a response was "correct" if S said the two figures covered the same amount of space (i.e., conservation of area in spite of perceptual differences.)

Volume: Several different aspects of the conservation of volume were studied by Lovell and Ogilvie (1961). The materials consisted of 25 plastic dice, a one gallon can, a one pint can, and water. To examine conservation of "internal" volume, two sets of 11 dice were built in identical 2x2x3 blocks. The E asked S, "If we made two boxes, one for each block, would there be as much room in one box as in the other? Why?" One block was then rearranged as 1x2x6 and S was asked, "If we make another box for this block, would it have the same amount of room in it as the previous box? Why?"

In studying conservation of "occupied" volume, E filled the pint can with water. Before filling the 1 gallon can, E put in a 2x3x2 block of dice. E asked the S, "If we now fill this can to the top do we still get the same amount of water as before, or do the dice make a difference? Why?' The 2x3x2 block was removed and E asked, "If we put this block (1x1x6) into the can, can we get as much water into the can now as we could with the previous block?" Conservation was shown if the S ascertained that the amount of water that could be poured into the can was independent of the arrangement of the dice as long as the number of dice remained the same.

Conservation of "displacement" volume was investigated by E saying, "Pretend that the gallon can is full of water and we place the block into the gallon can. Is it possible to put this block (2x3x2) in the gallon can without spilling water?" Those answering correctly were asked: "What happens if we put the 1x2x6 block in the can instead of the other block? What do you know about the water that spills over?" In addition, each S was asked whether any water would spill over if just one dye was lowered into (a) the full pint can, and (b) the full gallon can. If both answers were correct, S was asked whether the amount spilled would be the same. Those who answered affirmatively were asked to compare the amounts of water spilled by a cube made of plastic versus a cube of lead (of the same size) lowered into the full can.

Goodnow (1962) introduced a variation of the conservation of volume tasks. Two jars with equal volumes of water and two balls of clay, identical in size and shape were used. One ball was placed in each jar, and S was asked whether the volume remained the same. Then one ball was squashed, and the S was questioned again.

Frank (reported in Bruner, 1964) approached conservation of volume through the water jar problem. In the original problem, two jars of the same size and shape were filled with equal amounts of water. E asked if they were the same. The water in one jar was then poured into a second jar of a different size and S was asked if there was still the same amount of water in each jar. Ss who realized the volume of water was the same regardless of the size or shape of the jar had attained conservation of volume. Frank modified this problem by covering the different size jar with a screen while the water was poured. E then asked S if the amount of water was the same. Conservation occurred more readily in this case than in the former. The screen was then removed so that the Ss could see the new jar. Under the exposure condition many Ss who had formerly appeared to show conservation now responded to the perceptual cue of the jar's shape by reversing the original response.

Number: Dodwell (1960) studied the relation of perceived size to conservation of number. The materials were beads, two identical beakers, and two dissimilar beakers. First, S was asked to put 8 beads in each of a pair of identical beakers. Next, he was asked to put the same beads (recounting them) in dissimilar beakers. Finally, S was asked whether there was the same number of beads in each beaker. Conservation of number was shown by a "yes" response. If S said "no," he was asked which had more and why. Dodwell provides a complete list of 54 standard questions.

Wohlwill and Lowe (1962) investigated "number production," "equivalence" and "number versus length." The first two tasks did measure conservation, but were used to determine understanding of the concept of number. S was simply asked to give E six poker chips from a pile of poker chips, and then to match a row of seven poker chips that E had laid out. Conservation of number was shown by the number versus length tasks: (1) two parallel rows of 7 chips (one blue and one row red), were laid in equal rows. S was asked which had more chips. (This same question was asked for all the following comparisons.) (2) E extended the red row in both directions to make it twice as long.

(3) The red row was subdivided into 2 rows of 4 and 3 chips parallel to the S's blue row. (4) The red chips were put in a stack in front of the blue row. (5) The red chips were inserted into an opaque tube. (6) E laid 6 chips, but in a longer row than the 7 chip row before S. (7) All tasks were repeated for different number of chips.

Beilin (1965) modified the Wohwill-Lowe number conservation task. The apparatus consisted of two parts, a number rack for displaying three numbers, and a part which consisted of an expanding and contracting apparatus on which was placed a row of corks. The corks could be moved together or apart by pushing a handle at the end of the device. In the simplest task, number conservation was measured by requiring S to choose the number (of the three in the display rack) corresponding to the number of corks on the apparatus. After E expanded or contracted the array of corks, Ss were asked to make a second choice, this choice served as the measure of conservation of number. Nonconserving children chose a different number when the row was stretched or contracted.

Cork sheet surrounded the movable section so that a row of corks could be arranged on each side of the moving corks. (The corks in the center row were molded to the apparatus, while the rows of corks on the side were secured to the base by pins.) Pre-conservation tasks assessed the S's knowledge of "number production," "number equivalence," "equality," and "inequality." (1) Number production: S was shown a pile of thirteen red chips and told, "Give me six of these. Now give me eight. Now, give me twelve." (2) Number equivalence: E laid out a pile of seven red chips. S was told, "Make a pile with as many of your chips over here as these here." (3) Equality: E laid out six chips before himself and six chips before S and asked, "Do you have more, less or the same chips as I?" (4) Inequality: E laid out a pile of seven red chips before himself and five before S. "Who has more chips, you or I?" Then, "Who has less chips, you or I?"

Following these preliminary tests, the Ss were tested for conservation of number. There were 2 practice and 12 test trials, with each trial consisting of two parts. In the first part, S was shown the number apparatus with its three parallel columns of corks. One column was equal in number as well as length to the middle (stimulus) column. The other column was unequal in number and length to the middle one. S was instructed to choose the row which was "like" the middle one by pressing a button at the base of either of the two columns. If correct, he heard a buzzer and was given a token. After S responded, E expanded or contracted the stimulus column so that the first and last corks were aligned with the first and last corks of either the shorter or longer response column. No corks were removed or added. All contractions and expansions were made in sight of S. After each change, S was again asked to choose the column that was "like" the middle one, and a correct response was reinforced in the same manner. It was expected that the reinforcement would provide information as to which of the concepts (length covered vs. number) represented in the array was the one sought in the test. On half the trials, the incorrect column was shorter than the middle one, and, on the other half, it was longer. The number combinations changed in each trial. At the end of the 12th test trial, S was asked to explain his choice.

Conception of Space

Only one type of task relating to the representation of space was encountered in the literature on training effects. This task demonstrates a relatively advanced level of development at which children behave as if they perceived objects as located in a Euclidian space of horizontal and vertical coordinates.

Piaget and Inhelder (1956) reported on a study of the representation of water level. Materials consisted of straight and round-sided jars partially filled with water. The jars were covered and rotated to different angles. The task was for the Ss to anticipate the water level of the jar, tilted at various angles, and to copy the level with a line drawing on outline figures of the jars. A "Euclidian" representation of water level was achieved when the child drew horizontal lines, taking into account the generalized spatial surround, rather than lines perpendicular to the sides of the tilted jar.

Beilin, Kagan, and Rabinowitz (1966) investigated the influence of language and perceptual experience upon anticipation imagery in water level representation. The materials were round-bottomed Florence flasks, straight-sided jars, water, and outline drawings of the jars at different angles. The E placed a straight-sided gallon jar, half filled with red-colored water, on the desk before the Ss and covered the jar with opaque stretch material, thus hiding the water, while leaving the visual outline of the jar intact. When the jar was tilted, Ss were asked to choose from among eight pictures of jars the one representing the water level correctly. Each pictured jar was about 1 inch high and outlined in black. The water surface was indicated by a red line, and the water itself (either above or below the water line) was shown by red strippling. All jars on a page were tilted in the same direction as the jar used by E.

The various alternatives reflected the combination of two variables. The first involved the relation of the water line to a reference system either internal to the jar (parallel to base or parallel to side) or external to the jar (horizontal or diagonal). The second variation was of the body of water in the jar relative to the effect of gravity---that is, either the water was above the water line, defying gravity, or below, in accord with it. For each internal or external reference choice, there were two gravity options for a total of eight alternatives. The correct choice in each series, of course, was the "water level horizontal--water below line" representation.

Movement

Piaget investigated various operational structures related to understanding movement or the displacement of objects with respect to an ordered set of fixed positions (Flavell, 1963). Only one experiment dealing with the understanding of _spatial order_ was encountered in the survey of literature on training effects.

Greco, (see Smedslund, 1961a) studied the concept of direct and inverse spatial order. Materials were a cardboard tube and a wooden rod to which were fastened a black, a white, and a red bead, in that order. The rod with the beads was moved into the tube until completely hidden, and S was asked to tell which color would come out first at the other end. To determine whether S could maintain an accurate conception of order in spite of spatial rotation, the tube, with the rod and beads inside, was rotated slowly and horizontally 180 degrees. The question was again posed as to which bead would come out first. The number of rotations could be varied to make the task more difficult.

Transitivity

In the Piagetian system, the term "transitivity" refers to a logical operation involving the generalization of relationships. This approximates the meaning customarily assigned to the term. Flavell (1963) reports that while Piaget viewed the transitivity operation as accompanying the development of each conservation concept, later work has failed to confirm this hypothesis. Transitivity is obviously very useful as a cognitive tool in permitting the generation of relationships without direct experience with the particular quantities involved. In higher mental processes, what is essentially transitivity is often referred to as reasoning by analogy, and, as such, plays a significant role in problem solving and creativity.

Length: Braine (1959) employed a non-verbal technique for demonstrating transitivity of length. The materials were wooden sticks of various lengths and pieces of candy. S was to find the candy hidden under the longer (or shorter) of any two pieces of wood. Having learned this relative length cue, the Ss were shown that an upright piece of wood, A, was longer than a measuring stick, B, which in turn was longer than upright C. Transitivity was shown if the S could infer that the candy would be under the longer stick, A, when it was paired for the first time with the shorter stick, C.

Smedslund (1963) studied the development of "concrete" transitivity of length. Seven sticks of different lengths (six black and one blue) were laid on arms of cardboard which formed a V at the top and bottom of each, producing the Muller-Lyer Illusion. In this case, the V-forms were always placed so that the illusion would operate to make the shorter form appear longer. Two sticks were used in each trial, and the Ss were asked to judge which was longer, disregarding the V-shaped extensions. In later trials, the blue stick B (without the V-shaped arms) was placed between the pair AC, first close to C, and then moved near A. Last, B was removed and S was asked which of the A-C pair was longer. Ss with transitivity reported A longer than C on the basis of the previous comparison with B rather than on the direct comparison of A and C under conditions promoting the illusion. Ss without transitivity changed their answers to fit the perceptual illusion.

Weight: Smedslund (1961b) also studied transitivity of weight. Three objects made of plasticine (all of equal weight) and a balance were given to the S. He was asked to place two of the objects (e.g., red ball and green "sausage") on the balance to show that they were equal in weight. Then, the green sausage was balanced against a yellow "cake". Finally, E asked S if one of two previously unpaired objects (e.g., ball and cake) weighed more, the same, or less than the other. An understanding of transitivity would, of course, permit the reasoning, if A equals B, and B equals C, then A also equals C.

Preference: Smedslund (1960) first determined each S's preferences by presenting four sets of pictures of different objects. There were three pictures in each set, and S was asked to indicate the one he preferred. In the test for transitivity, each set of three pictures was again shown to S. A new card was produced picturing two objects from the set (e.g., a doll and a rabbit) with a boy (or girl) standing between them, pointing at one of the objects. The S was asked to tell which object the boy (or girl) the preference card preferred. This first preference card was left on the table along with the three pictures and a second was produced with the figure pointing at one of a second pair (e.g., doll and teddy-bear). E then summarized verbally the meaning of the first two preference cards. A third card was presented showing the boy standing between the last two objects of the set to be compared (e.g., rabbit and teddy-bear) but not pointing at either. S was to tell which of the two objects the boy (or girl) preferred. A correct response would, of course, imply transitivity of the relationships established by the first two cards.

Ss were tested on all four sets of pictures.

Logical Groupings

A considerable portion of Piaget's observation and theory dealing with the concrete level is concerned with the nature of the operations by which children establish logical classes and relationships. Flavell (1963) emphasizes that Piaget's approach has been logical rather than empirical in that his theory accounts for the various conceptual organizations, but makes no claim that children actually manipulate the formal postulates of logic which compose it. Since there are nine grouping structures in the theory, each cannot be presented here. As an example, the first of the studies summarized below can be categorized under Piaget's "Grouping I." In general terms, this operation permits children to see subclasses as combined into a superordinate class, and the latter as divisible into subclasses. The subjects in the study were mainly unable to conceive of objects belonging simultaneously to a superordinate class and a sub-class.

Smedslund (1960) showed how a complex grouping of objects demonstrated the consequences of the absence of reversibility at the concrete level. S was given three white wooden beads, 17 brown wooden beads, and several uncolored glass beads and asked to tell whether there were more wooden beads or brown beads. While concrete Ss were able to think alternatively in terms of one or the other of the classifications: (e.g., "brown" or "wooden"), they could not apply both concepts simultaneously to the same set of objects.

Weinberg (1963) used a somewhat related task from the Weigl Color-Form Test to classify the grouping strategies of children of various ages. Twelve pieces of tile (three shapes and four colors) were provided and S was told to group them in some way. Then, S was asked to regroup in a different way. Groupings by both form and color were classified at the highest level, followed by form alone, by color alone, or by neither.

Inhelder and Piaget (1958) investigated the relation between age and a grouping task. Ss were provided with 5 or 6 cups containing various colored tokens and were asked to make up all possible pairs of tokens. At the concrete level, the Ss' combinations remained incomplete and were generated by unsystematic strategies such as trial and error. At the level of formal operations, the Ss used a systematic method, employing the operations of permutations and combinations (without explicit formulation of mathematical expressions, but with a performance based on an exhaustive method.) Goodnow (1962) used a similar method in a cross-cultural investigation.

Bruner and Olver (1963) conducted a cross-sectional study of the development of conceptual grouping at the verbal level. S was given two words, such as banana and peach, and asked to give a concept that would include both objects. S was then given a third word, such as potato, and asked to expand the concept so as to include all three words. This procedure was continued with successively more discrepant concepts (the particular list in the above example ended with the word "stones") so that the organizing concept had to become more and more generalized (e.g., abstract). True conceptual groupings on the basis of one or more common attributes were classified as _superordinate_, and were found to increase with age. _Complex_ groupings were of several varieties, all involving a failure to subordinate the entire list to any single, common attribute. _Thematic_ groupings involved telling stories involving the objects. An analysis of the language used by the Ss showed that perceptual attributes (a) tended to be associated with complex groupings, and (b) to decline with age. Use of functional (as opposed to perceptual) attributes (a) increased with age and (b) at any age level was associated with superordinate grouping strategies.

These observations were extended by Bruner (1964) in a paper analyzing grouping strategies both from the point of view of the attributes used (e.g., perceptual vs. functional features) and in terms of the logical structure of the groups (e.g., complexes without a uniform grouping rule and superordinate concepts with a systematic rule applicable to all instances). An interesting conceptual experiment was discussed in this latter paper which clearly relates grouping strategy to the formal modes of problem solution to be discussed in Appendix B. Children in a game situation are asked to develop hypotheses about vague or fuzzy pictures that are gradually brought into focus. Younger children tend to pose highly specific questions or hypotheses, while older children ask questions that are general enough to eliminate (or suggest) whole classes of specific hypotheses.

Verbal Comprehension

The study by Case and Collinson (1962) is not classifiable under any of the above categories, since the test task utilized written material and provided scores at three levels of conceptual development. While interesting, this procedure is related to Piaget's theory only in a very loose and general way. This research has already been described in detail on page 34 of the text.

APPENDIX B

A Summary of Problem Tasks Assigned Primarily to the Level of Formal Operations[1]

---

[1] From Inhelder and Piaget (1958).

Appendix B

A Summary of Problem Tasks Assigned Primarily to the Level of Formal Operations

Equality of Angles of Incidence and Reflection

Apparatus: Essentially a type of pin ball game. Balls are launched with a tubular spring device that can be pinated in an arc around a fixed point. The ball is shot against a projection wall and thus rebounded to the interior of the table, where targets are placed at different points from trial to trial.

Task: The S is instructed to aim the projection device so as to hit the target with the ball. Since it is impossible to aim directly at the target, the S obviously has to aim at a point on the wall such that a hit will be scored on the rebound.

Concept: The principle constituting a formal solution of the problem is of course that the angle of reflection of the ball off the projection wall is equal to its angle of incidence. As far as accuracy is concerned, however, younger Ss at the level of concrete operations can perform the task quite effectively on the basis of an empirical, trial and error strategy without ever generating the formal solution. Inhelder and Piaget, who were especially interested in this task, maintained that it provides an unusual opportunity to isolate operational mechanisms involved in formal reasoning, where the reasoning process deals with the manipulation of notions that have already been acquired.

Floating Bodies

Apparatus: a number of disparate objects (some float and some do not) and several buckets of water.

Task: The S is asked to classify the objects according to whether or not they will float on water and to explain the basis of each of his classifications. The S then experiments with the objects to see which ones will float. Finally, S is asked to summarize his observations in the form of a law. After the classification, older Ss are given three cubes of equal volume, but having different densities, and an empty cube with clear plastic walls and density of about unity to facilitate accurate comparisons with the density of water.

Concept: Objects float if their density of specific gravity is less than that of water. In this problem two relationships were essential to the solution: density (i.e., the relation of weight to volume) and specific gravity (i.e., the relationship between the weight of the object and an equivalent volume of water). The concept differs from the "Equality of Angles" problem in that it cannot be derived from concepts entirely accessible at the level of concrete operations. Even though some of the simpler contradictions can be overcome by means of concrete operations, the elimination of more subtle ones and the formulation of a unified explanation requires the use of implications (i.e., the intervention of formal propositional operations).

Rod Flexibility

Apparatus: A large basin of water and a set of rods differing in composition (e.g., steel, brass, etc.), length, thickness, and cross-section form (i.e., round, square, rectangular). Three different weights can be screwed on to the ends of the rods, which, in turn, can be shortened or lengthened by varying the point where they are clamped to a vertical support. The rods extend horizontally over the water basin with the weights placed so as to exert a force perpendicular to the surface of the water. Maximum rod flexibility is indicated when the end of the rod touches the water.

Task: S determines whether or not the rod is flexible enough to reach the water level. His methods are observed and comments on the variables believed to influence flexibility are noted. S is eventually asked for a proof of his assertions as to the latter.

Concept: The flexibility of a rod depends on its composition length, thickness, and the form of its cross-section. All other things being equal, flexibility varies as a function of the weight placed at the end of the rod. The concepts involved in the solution of this problem are no more complex than the previous problems. However, there is greater empirical difficulty since a complete solution requires that five distinct variables be varied systematically and independently. Thus, the aspects of formal reasoning demonstrated by this problem involve the differentiation of relevant variables and classification of systematic effects.

Pendulum Oscillation

Apparatus: A pendulum in the form of an object suspended from a string. S is given the means to vary the length of the string, the weight of the suspended objects, the amplitude, etc.

Task: The subject is to find what determines the frequency of oscillation.

Concept: The frequency of the oscillations of a pendulum is related solely to the length of the string. In the previous problems, S had to separate out the factors in order to determine their respective effects in a multivariate situation. In the present task there are many variables to be manipulated, but only one (length of string) actually plays a causal role. Thus, all must first be isolated, varied, and then excluded in order to find the causal factor. The formal process exemplified here is the operation of "exclusion" that is, eliminating apparently relevant variables (e.g., weight of object, height of drop point, etc.) in order to isolate the solution to the problem.

Falling Bodies on an Inclined Plane

Apparatus: A ramp mounted on a board that can be adjusted to various angles on inclination. Balls of various sizes are rolled down the ramp from several release points, a curved section at the bottom, and are lofted into the air in a parabolic curve. A box with eight compartments extends beyond the curved end of the ramp, and each ball lands in one of these boxes as a means of recording the length of the bound.

Task: The S is asked to find the relationship between the above variables and the distance the ball travels.

Concept: The length of the bound varies only as a function of the height of the release-point on the plane, being unaffected by slope as long as height remains constant. The factors of mass or weight of the ball must be excluded to solve the problem. This problem demonstrates the disjunctive operations of formal thought. In this case "disjunction" means that either the distance or the slope or both (i.e., height) are the determining factors. Thus the S sees that, even though a given slope remains constant, an increase in height brings about a simultaneous increase in distance. If, on the other hand, the height remains constant, increasing slope decreases the distance. The authors emphasize that the solution of this problem in part depends on the way the variables are presented to the S.

The Role of Invisible Magnetization

Apparatus: A large round board divided into eight sectors of different colors and equal surface area, with opposite sectors matching in color. A metal bar is attached to a nonmetallic, rotating disk at the center of the board. Boxes, each with different symbols (opposite pairs the same symbol), are placed in the eight sectors of the board. The rotating disk is spun and eventually stops with the metal bar pointing to one pair of boxes. It turns out that the bar always points to the _same_ pair of boxes. While the boxes are moved to different sectors, they are always placed with the members of a pair opposite each other. The boxes are also unequal in weight, providing another variable. As might be guessed, the pair of boxes to which the bar inevitably points contain several magnets concealed in wax.

Task: The S determines why the metal bar always ends by pointing at the same pair of boxes.

Concept: The metal bar will always point to the boxes containing the magnets. The significant aspect of this task is that the S's cannot arrive at this concept directly. In order to arrive at a solution to the problem, each variable must be systematically varied and excluded, with the correct solution calling for the formulation of an hypothesis _beyond_ the given variables. In sum, the present problem demonstrated that transition from concrete to formal operations is distinguished by the appearance of a complete combinatorial system whose various types of disjunction and exclusion are continuously linked to implications.

Combinations of Colored and Colorless Chemical Bodies

Apparatus: Four similar flasks containing perceptually identical colorless and odorless liquids. Flask #1 contains diluted sulphuric acid, flask #2 water, #3 oxygenated water, and #4 thiosulphate. S's are also given a bottle (with an eye dropper) containing potassium iodide (call it "g"). The oxygenated water, when combined with oxidized potassium iodide in an acid medium, yielded a yellow color (1+3+g). The water in flask 2 is neutral, so adding it will not change the color. The thiosulphate in flask 4 will bleach the mixture.

Task: E presents S with two glasses, one containing 1 + 3, the other containing only plain water. In view of the S, E pours several drops of g into the two glasses and notes the different reactions. The S is asked to reproduce the color yellow, using flasks 1, 2, 3, 4, and g as he wishes.

Concept: This problem involved combinations directly in that there are elements whose combinations are indispensable if the demonstration is to be reproduced. A combinatorial system is implicit in the S's potential ability to link a set of elementary associations or correspondences with each other in all possible ways so as to draw from them the relationships of implication, disjunction, and exclusion and thereby solve the problem. In previous problems the variables could be disassociated and combined at will or simply made to correspond without going beyond the level of more or less random trial and error. The present task requires systematic combination if valid results are to be obtained.

Conservation of Motion in a Horizontal Plane

Apparatus: A spring device launches balls of varying diameters and weights which in turn roll along a horizontal plane until stopping.

Task: S's are asked to predict the stopping points of the balls while varying their size and weight. Later the S's are asked to explain the basis on which their predictions were made.

Concept: If no external obstacle were to interfere, an object would maintain a uniform rectilinear motion (principle of inertia). Actually, a number of factors prevent the free operation of inertia (e.g., friction due to weight, air resistance due to volume, etc.). As a result, two problems arise which must be resolved by formal thought: (1) the problem of what is theoretically possible (or how does one come to understand the conservation of motion by inertia given that it is never observable?), and (2) the problem of the relatively possible. This problem is interesting because concrete operations of serial ordering and correspondence formation allow the establishment of accurate empirical relationships between the properties of the balls and their stopping points. Yet, the idea of conservation of movement by inertia transcends the realm of the "concrete", for such conservation cannot actually be achieved under the experimental conditions available to S. Thus, the formal operation of implication must be achieved in order to formulate the "possible."

Communicating Vessels

Apparatus: Two vessels of differing shape and volume, both containing water, are mounted on vertical rods. The vessels are connected to each other at the bottom by a loop of flexible tubing (hence, "communicating"). The vessels can be raised or lowered.

Task: The authors do not state the exact nature of the task as assigned to the S, but it appears that S is to explain the observed equilibrium between the two levels of water when the vessels are raised or lowered.

Concept: In every equilibrium two forms of reversibility operate simultaneously: inversion, which corresponds to the additions or eliminations effected in the parts of the system that come into equilibrium, and reciprocity, corresponding to the symmetries or compensations between these parts. In the problem of communicating vessels, reciprocity serves to express the compensatory actions between separate vessels, transformations by inversion express the rise and fall of the water level. Changes in water level are brought about not by adding or taking away water, but by raising and lowering the vessels. The fundamental concept to be discovered is that the pressure exerted by the liquid in one flask is at the same time equal to, and acting in the opposite direction from, the pressure exerted by the liquid in the other flask, thus establishing equilibrium. Since the vessels have neither the same shape nor the same volume, one has to exclude these two factors in finding the law. Air pressure can be disregarded, being it is equivalent for the two columns of water. Certain rough intuitions of equilibrium occur at a very early age, but the concept is not really understood before the formal level, when the S can both distinguish and coordinate inversions, reciprocities, and correlativities. A major drawback of the problem is that the pressure intrinsic to the liquid is usually overlooked by the unsophisticated. Thus, detailed explanations are not given until knowledge is acquired through formal training in science.

Equilibrium in the Hydraulic Press

Apparatus: Two "communicating vessels" of different sizes and shapes, mounted on a pair of vertical stands, and connected at the base by a loop of flexible tubing. Vessel A is provided with a piston that can be loaded with varying weights, and the amount of pressure exerted on the liquid by the piston is varied by adding the weights. Various liquids are used as an additional factor (e.g., water, alcohol, or glycerine).

Task: The S is to observe and explain the effects of different weights and liquids on the change in the height of the liquid in Vessel B.

Concept: The height of the liquid in Vessel B is directly related to the weight of the piston (since the pressure exerted on the liquid in A is directly proportional to the weight) and inversely a function of the density of the liquids. The reaction of the fluid to the weight of the piston and the resistance reaction of the different liquids can be made tangible by varying the weight of the piston and the density of the liquid. Thus, this problem involves equilibrium plus the transmission of forces. The problem is to understand that the force exerted by the piston is transmitted in a uniform manner through the entire liquid and that the equilibrium between action and reaction relates not only to the surface of the liquid but to the entire system. In this particular case the transition from concrete to formal thought involves moving an explanation based on simple observed correspondences between the weights and the displacements of the liquid to an explanation expressing the complete transmission of force as a function of weight and inversely as a function of density.

Equilibrium in the Balance

Apparatus: Two forms of a balance scale: (a) a conventional balance with varying weights (standard types) that can be hung at different points along the balance cross-bar, and (b), a balance equipped with baskets that can be moved along the cross-bar. Dolls were used as weights on the second balance.

Task: The S is asked to vary the weights and distances from the fulcrum on the crossbar in such a way as to obtain equilibrium, and then to explain the relationship of weight and distance to equilibrium.

Concept: The main purpose of the problem is to find out why the notion of proportions does not appear until around the age of 13 or 14 years. Finding the law presupposed the construction of the proportion $W/W' = L'/L$ where W and W' were the weights and L and L' were the relative distances of the weights from the fulcrum. Spelling out this explanation implied an understanding of the proportion $W/W' = H'/H$ (H and H' are the respective heights of the weights from the base).

Hauling Weight on an Inclined Plane

Apparatus: A toy dumping wagon is hauled up an inclined rail of variable inclination by counterweights suspended at the opposite end of a cable. The three main variables were the angle of the rail, the number of counterweights, and the weight of the wagon (additional weights could be put in the wagon).

Task: S predicts the movements or equilibrium position of the wagon as a function of the three variables (i.e., weight in wagon, counterweights suspended at the end of the cable, and the inclination of the rail). Inclination was calculated not in terms of its angle measured in degrees, but in terms of the height (h).

Concept: The law of equilibrium to be found is W/M = h/H (or counterweight/wagon = height of inclination/constant length of rail). This problem is similar to the preceeding "Equilibrium in Balance" problem, but is especially designed to bring out physical work relationships. More work is required to pull a given weight along a steeper rail than along a more gently sloping rail. As long as the S is limited to using concrete operations of classes and relations, he cannot determine the law. With formal thought, S's use a remarkably different approach to the problem. Instead of getting lost in the inventory of actual cases, the S very quickly turns to a selection of crucial cases and seeks to coordinate the three factors into a single law.

The Projection of Shadows

Apparatus: The principle materials are a baseboard with a screen attached to one end and a light source at the other end, and four mounted rings of varying diameters. The rings are placed between the light source and the screen, so as to cast a shadow on the latter. Both light source and the rings can be moved along the baseboard.

Task: The S is asked to find two shadows that cover each other exactly, using two unequal rings. S thus has to vary the size of rings and the distance in attempting to project shadows of equal size.

Concept: Rings placed between a light source and a screen cast shadows of a size directly proportional to their diameters and inversely proportional to the distance from the light source. The equilibrium schema of the two previous problems called for finding proportions derived from a model of physical equilibrium. The proportions in connection with the projection of shadows differ in that they are of an essentially geometrical nature denoting relationships between physical distances and diameters that can be explained in terms of simple projective geometry.

Centrifugal Force and Compensations

Apparatus: Three metal balls of different sizes and weights are placed on a disc at three different distances from its center. The disc is rotated at a constant speed until centrifugal force causes the balls to roll off.

Task: S is asked to predict the order in which the balls will leave their initial positions and to explain the basis of the prediction.

Concept: Since the speed of the disc is constant after the initial acceleration, the S need isolate only the factors of mass and radius---i.e., he need understand only the following two relationships: a ball is displaced sooner in direct proportion to its weight and later in inverse proportion to the distance from the center of the disc. Consequently, a problem of compensation arises---a heavy ball placed at a point nearer the center may move at the same time as a lighter one closer to the periphery (in fact, the three weights are calculated in such a way as to compensate exactly for the three distances). In this problem the authors attempted to define the relationship between the proportionality schema and the schema of multiplicative compensation. Increasing the value of one of the variables gives the same result as an increase or decrease in the value of the other. This is not the first problem requiring the use of the compensation schema. However, in the present case two possibilities are open to the S. He can construct proportions involving numerals (which he could not do in the "Flexibility of the Rod" problem), or he can isolate the factors that determine equilibrium in terms of the "all other things being equal" method (which he could not do in the "Traction" problem). The author's aim was to discover whether, psychologically, proportions carry with them the idea of compensation or whether it is the other way around.

Correlations

Apparatus: S's are shown 40 cards, each showing a human face drawn on it. Eyes and the hair are colored according to the following four combinations: (1) a = blue eyes and blond hair, (2) b = blue eyes and brown hair, (3) c = brown eyes and blond hair, and (4) d = brown eyes and brown hair.

Task: S is given a set number of cards and asked whether there is a relationship between eye color and hair color (i.e., not whether there is such a relationship in real life, but whether one can be discovered in the given data). S can easily count the number of cases that confirm and disconfirm a hypothesized relationship between the two variables.

Concept: The problem involves discovering that there is a mere correlation between eye color and hair color rather than a one-to-one relationship. E can proceed in two different ways: (1) let the S form his own classification or (2) give him the cards already classified according to the four possibilities. The latter method puts more emphasis on the possible numerical combinations, asking the S in each case to estimate the relevant relationships. In addition the S can be shown two different sets, and then asked which shows the clearest correlation. Finally, S can be asked to remove cards in such a way as to strengthen the correlation; then he can be asked to discuss which of the four associations he used as a basis for eliminating cards. Once the initial difficulties have been overcome, the problem is to discover that correlation involves a probabilistic orientation that takes more than one outcome into account.

References

Ausubel, D.P. and Ausubel, Pearl. "Cognitive development in adolescence." Review of Educational Research, 1966, 36, 403-413.

Beilin, H. and Franklin, I.C. "Logical operations in area and length measurement: Age and training effects." Child Development, 1962, 33, 607-618.

Beilin, H. "Learning and operational convergence in logical thought development." Journal of Experimental Child Psychology, 1965, 2, 317-339.

Beilin, H., Kagan, J.L., and Rabinowitz, R. "Effects of verbal and perceptual training on water level representation." Child Development, 1966, 33, 317-329.

Berlyne, D.E. Structure and Direction in Thinking. New York: John Wiley & Sons, 1965.

Braine, M. "The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods." Psychological Monographs, 1959, 73, Whole No. 475.

Bruner, J.S. "The course of cognitive growth." American Psychologist, 1964, 19, 1-15.

Bussis, Anne M., From theory to the classroom: Background information on the first grade project in New York City Schools. New York: Board of Education of the City of New York, 1965.

Case, D. and Collison, J.M. "The development of formal thinking in verbal comprehension." British Journal of Educational Psychology, 1962, 32, 103-111.

Clayton, Martha D. and Jackson, D. "Equivalence range, acquiescence, and overgeneralization." Educational and Psychological Measurement, 1961, 21, 371-382.

Cronbach, L.J. "Course improvement through evaluation." Teachers College Record, 64, 1963, 672-683.

Dodwell, P.C. "Children's understanding of number and related concepts." Canadian Journal of Psychology. 1960, 14, 191-205.

Dressel, P. L. and Mayhem, L. B. *General Education; Explorations in Evaluations*. Washington, D.C.: American Council on Education, 1954.

Elkind, D. "Quantity conceptions in college students." *Journal of Social Psychology*, 1962, 57, 459-465.

Flavell, J. H. *The Developmental Psychology of Jean Piaget*. New York: Van Nostrand, 1963.

Freyberg, P. S. "Concept development in Piagetian terms in relation to school attainment." *Journal of Educational Psychology*, 1966, 57, 164-168.

Goodnow, Jacqueline J., "A test of milieu effects with some of Piaget's tasks." *Psychological Monographs*, 1962, 76, No. 555.

Goodnow, Jacqueline J. and Bethon, Gloria. "Piaget's tasks: the effects of schooling and intelligence." *Child Development*, 1966, 37, 573-582.

Hunt, J. McV. *Intelligence and Experience*. New York: Ronald Press, 1961.

Husek, T. R. *Different kinds of evaluation and their implication for test development*. Paper presented at annual meeting of the American Educational Research Association, Chicago, 1966.

Huttenlocher, J. "Some effects of negative instances on the formation of simple concepts." *Psychological Reports*, 1962, *11*, 35-42.

Huttenlocher, J. "Development of formal reasoning on concept formation problems." *Child Development*, 1964, *35*, 1233-1242.

Inhelder, Barbel and Piaget, J. *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books, 1958.

Inhelder, Barbel, Bove, Magali, Sinclair, Hermine, and Smock, C. D., "On cognitive development." *American Psychologist*, 1966, 21, 160-164.

Jackson, D. N., Messick, S., and Myers, C. T. "The role of memory and color in group and individual embedded-figures measures of field-independence." *Educational and Psychological Measurement*, 1964, 24, 177-192.

Kagen, J., Moss, H. A. and Sigel, I. E. "Psychological significance of styles of conceptualization." *Monographs of the Society for Research in Child Development*, 1963, 28, 73-112.

Lovell, K. And Ogilvie, E. "A study of the conservation of substance in the junior school child." *British Journal of Educational Psychology*, 1960, 30, 109-118.

Lovell, K. and Ogilvie, E. "The growth of the concept of volume in junior school children." *Journal of Child Psychology and Psychiatry*, 1961, 2, 118-126.

Messick, S. and Kogan, N. "Differentiation and compartmentalization in object-sorting measures of categorizing style." *Perceptual and Motor Skills*, 1963, 16, 47-51.

Pettigrew, T. F. "The measurement and correlates of category width as a cognitive variable." *Journal of Personality*, 1958, 26, 532-544.

Piaget, J., and Inhelder, B. *The Child's Conception of Space*. London: Routledge & Paul, 1956.

Piaget, J., Inhelder, B., and Szeminska, A. *The Child's Conception of Geometry*. New York: Basic Books, 1960.

Pratoomraj, S. and Johnson, R. C. "Kinds of questions and types of conservation tasks as related to children's conservation responses." *Child Development*, 1966, 37, 343-353.

Scriven, M. *The Methodology of Evaluation*. Bloomington, Indiana: Indiana University, 1965. (mimeograph)

Smedslund, J. "Transitivity of preference patterns as seen by pre-school children." *Scandinavian Journal of Psychology*, 1960, 1, 49-54.

Smedslund, J. "The acquisition of conservation of substance and weight in children: I. Introduction." *Scandinavian Journal of Psychology*, 1961a, 2, 11-20.

Smedslund, J. "The acquisition of conservation of substance and weight in children: II. External reinforcement of conservation of weight and of the operations of addition and subtraction." *Scandinavian Journal of Psychology*, 1961b, 2, 71-84.

Smedslund, J. "The acquisition of conservation of substance and weight in children: III. Extinction of conservation of weight acquired "normally" and by means of empirical controls on a balance." Scandinavian Journal of Psychology, 1961c, 2, 85-87.

Smedslund, J. "The acquisition of conservation of substance and weight in children: IV. Attempt to extinction on the visual components of the weight concept." Scandinavian Journal of Psychology, 1961d, 2, 153-155.

Smedslund, J. "The acquisition of conservation of substance and weight in children: V. Practice in conflict situations without external reinforcement." Scandinavian Journal of Psychology, 1961e, 2, 156-160.

Smedslund, J. "Development of concrete transitivity of length in children." Child Development, 1963, 34, 389-405.

Uzgiris, I.C. "Situational generality of conservation." Child Development, 1964, 35, 831-841.

Weinberg, N. "Relativism, Self-centering and conceptual level." Child Development, 1963, 34, 443-450.

Witkin, H.A., Dyk, R.B., Faterson, H.F., Goodenough, D.R., and Karp, S.A. Psychological Differentiation. New York: John Wiley and Sons, Inc., 1962.

Wohlwill, J. and Lowe, R.C., "Experimental analysis of the development of the conservation of number." Child Development, 1962, 33, 153-167.